

Integer-Valued Autoregressive Processes with Periodic Structure

Magda Monteiro

Escola Superior de Tecnologia e Gestão de Águeda

Research Unit ‘Matemática e Aplicações’

Universidade de Aveiro

Portugal

Manuel G. Scotto* and Isabel Pereira

Departamento de Matemática

Research Unit ‘Matemática e Aplicações’

Universidade de Aveiro

Campo Universitário de Santiago

3810-193 Aveiro, Portugal

Abstract

In this paper the periodic integer-valued autoregressive model of order one with period T , driven by a periodic sequence of independent Poisson-distributed random variables, is studied in some detail. Basic probabilistic and statistical properties of this model are discussed. Moreover, parameter estimation is also addressed. Specifically, the methods of estimation under analysis are the method of moments, least squares-type and likelihood-based ones. Their performance is compared through a simulation study.

Keywords: Count processes; Periodic autocovariances; Binomial thinning

*Manuel G. Scotto, Departamento de Matemática, Universidade de Aveiro. Campo Universitário de Santiago, 3810-193 Aveiro, Portugal. Tel.: +351 234370670; fax: +351 234370670. Email: mscotto@ua.pt.

1 Introduction

Since their introduction by Bennett (1958) and Gladyshev (1961, 1963) much attention has been given to periodically correlated (or cyclostationary) processes, partially because of their wide applicability to hydrology (Vecchia 1985, Salas 1993, and Tesfaye et al. 2006), economics (Franses 1994, Franses and Paap 2004, and Haldrup et al. 2007), meteorology (Bloomfield et al. 1994, Lund et al. 1995, and Lund et al. 2007), and signal processing (Gardner et al. 2006). Further examples can be found in Hurd and Miamee (2007) and the references therein. A large part of the literature on this topic is devoted to the so-called periodic autoregressive moving average (PARMA) models which are extensions of the commonly used ARMA models, having parameters which vary periodically in time. The analysis of basic probabilistic properties of PARMA models as well as statistical inference and forecasting techniques, have been addressed by Lund et al. (2006), Shao (2006), Basawa et al. (2004), Shao and Ni (2004), Basawa and Lund (2001), Lund and Basawa (2000), and Bentarzi and Hallin (1994) among other authors.

It is worth to mention that all references given in the previous paragraph deal with the case of continuous-valued (i.e. conventional) periodically correlated processes. In contrast, however, the analysis of periodically correlated series of counts has not received much attention in the literature. This paper aims to give a contribution towards this direction. Motivation to include discrete data models comes from the need to account for the discrete nature of certain data sets, often counts of events, objects or individuals. Examples of applications can be found in the analysis of the number of rainy days (Cui and Lund, 2009), time series of count data that are generated from stock transactions (Quoreshi, 2006) where each transaction refers to a trade between a buyer and a seller in a volume of stocks for a given price, statistical control process (Weiß, 2009), telecommunications (Weiß, 2008), and also in the analysis of optimal alarm systems (Monteiro et al., 2008), experimental biology (Zhou and Basawa, 2005), social science (McCabe and Martin, 2005), and queueing systems (Ahn et al. 2000).

In this paper, we investigate basic probabilistic and statistical properties of the Periodic INteger-valued AutoRegressive process of order one with period T (hereafter PINAR(1) $_T$) defined by

the recursive equation

$$X_t = \phi_t \circ X_{t-1} + Z_t, t \in \mathbb{N} \quad (1)$$

with $\phi_t = \alpha_j \in (0, 1)$ for $t = j + kT, (j = 1, \dots, T, k \in \mathbb{N}_0)$, where the *thinning* operator \circ is defined as

$$\phi_t \circ X_{t-1} \stackrel{d}{=} \sum_{i=1}^{X_{t-1}} U_{i,t}(\phi_t),$$

being $(U_{i,t}(\phi_t))$, for $i = 1, 2, \dots$, a periodic sequence of independent Bernoulli random variables (r.v's) with success probability $P(U_{i,t}(\phi_t) = 1) = \phi_t$. Furthermore, it is assumed that (Z_t) constitutes a periodic sequence of independent Poisson-distributed random variables with mean v_t ($Z_t \sim P(v_t)$) with $v_t = \lambda_j$ for $t = j + kT, (j = 1, \dots, T, k \in \mathbb{N}_0)$, which are assumed to be independent of X_{t-1} and $\phi_t \circ X_{t-1}$. To avoid ambiguity, T is taken as the smallest positive integer satisfying (1). It is important to stress the fact that discreteness of the process (X_t) is ensured by the \circ -operator since this operator incorporates the discrete nature of the variates and acts as the analogue of the standard multiplication used in the continuous-valued moving average model. The concept of thinning is well known in classical probability theory and has been in use in the Bienaymé-Galton-Watson branching processes literature as well as in the theory of stopped-sum distributions. Potential applications of INAR processes with periodic structure can be found in the analysis of demographic or labor market data set. As an illustrative example we display in Figure 1 the time series trace plot of the monthly number of short-term unemployed people in Penamacor County (Portugal), from January 1997 to December of 2007.

(Figure 1 about here)

Figure 2 displays the sample means, the sample variances by months of the year, and the autocorrelation function.

(Figure 2 about here)

Further examples can be found in the analysis of international tourism demand (Brännäs et al. 2002 and Brännäs and Nordström 2006). In general, the data exhibits a strong from of periodic variation over the day of the week in addition to a strong seasonal variation over the year, with a peak in July-August.

The rest of the paper is organized as follows: in Section 2, we demonstrate the existence of an almost surely unique non-negative integer-valued periodically stationary process satisfying (1). Expressions for the periodic mean and autocovariance of the periodically stationary distribution are also given. Parameter estimation is covered in Section 3. In Section 4 the results are illustrated through a simulation study. Finally, some concluding remarks are given in Section 5.

2 Basic properties of the PINAR model

The analysis of the existence and uniqueness of a periodically stationary and causal PINAR(1)_T process can be obtained via the analysis of the multivariate integer-valued autoregressive process introduced by Latour (1997). Note that equation (1) admits the representation

$$\mathbf{Y}_t = \mathbf{A} \circ \mathbf{Y}_{t-1} + \boldsymbol{\zeta}_t \quad (2)$$

with $\mathbf{Y}_t = (X_{1+tT}, X_{2+tT}, \dots, X_{T+tT})'$, where $'$ denotes matrix transpose,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \dots & \alpha_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \prod_{k=0}^{T-1} \alpha_{T-k} \end{pmatrix}$$

and

$$\boldsymbol{\zeta}_t = \mathbf{B} \circ \mathbf{Z}_t = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \alpha_2 & 1 & 0 & \dots & 0 & 0 \\ \alpha_3 \alpha_2 & \alpha_3 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \prod_{j=0}^{T-3} \alpha_{T-1-j} & \prod_{j=0}^{T-4} \alpha_{T-1-j} & \prod_{j=0}^{T-5} \alpha_{T-1-j} & \dots & 1 & 0 \\ \prod_{j=0}^{T-2} \alpha_{T-j} & \prod_{j=0}^{T-3} \alpha_{T-j} & \prod_{j=0}^{T-4} \alpha_{T-j} & \dots & \alpha_T & 1 \end{pmatrix} \circ \begin{pmatrix} Z_{1+tT} \\ Z_{2+tT} \\ \vdots \\ Z_{T+tT} \end{pmatrix}$$

with $(\boldsymbol{\zeta}_t)$ being a sequence of i.i.d. integer-valued random vectors independent of the operators, with finite mean given by $E(\boldsymbol{\zeta}_t) = \mathbf{B}(\lambda_1, \lambda_2, \dots, \lambda_T)'$ and $V(\boldsymbol{\zeta}_t) = \mathbf{B} \sum_{\mathbf{Z}} \mathbf{B}'$ where $\sum_{\mathbf{Z}} = Cov(\mathbf{Z}_t, \mathbf{Z}_{t+j}) = \text{diag}(\lambda_1, \dots, \lambda_T)$. In view of the fact that the eigenvalues of the matrix \mathbf{A} are lesser than one and that $\boldsymbol{\zeta}_t$ is independent of \mathbf{Y}_s ($s < t$), it follows by Proposition 3.1 in Latour

(1997, p. 236) that there exist an almost surely unique non-negative integer-valued stationary process satisfying (2). Next we obtain the periodically stationary distribution of (X_t) . First, however, we prove the following result. For simplicity in notation we define

$$\beta_{t,i} = \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j} & i > 0 \\ 1 & i = 0 \end{cases},$$

which can be rewritten as

$$\beta_{t,i} = \begin{cases} \beta_{t,j} \beta_{T,T}^k, & i = j + kT; j = 1, 2, \dots, T \\ 1, & i = 0 \end{cases}.$$

Note that the sequence $(\beta_{t,i})$ is periodic with period T .

Proposition 2.1. *For a fixed value of $j = 1, \dots, T$, with $T \in \mathbb{N}$, the process (X_t) for $t = j + kT$ and $k \in \mathbb{N}_0$ is an irreducible, aperiodic and positive recurrent (and hence ergodic) Markov chain. Moreover, the stationary distribution of (X_t) is given by that of*

$$V_j = \sum_{m=1}^{+\infty} \sum_{a=0}^{T-1} \left(\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1} \right) \circ Z_{T(m+1)-a} + \sum_{m=0}^{j-1} \beta_{j,m} \circ Z_{j-m},$$

where the series converges almost surely and also in L_2 .

Proof. See Appendix A. □

Now we are prepared to obtain the periodic mean and autocovariance function of (X_t) .

Lemma 2.1. *For a fixed value of $j = 1, \dots, T$, with $T \in \mathbb{N}$, $t = j + kT$ and $k \in \mathbb{N}_0$*

$$\mu_j = \mu_t = E(X_t) = V(X_t) = \frac{\sum_{k=0}^{j-1} \beta_{j,k} \lambda_{j-k} + \beta_{j,j} \sum_{i=0}^{T-j-1} \beta_{T,i} \lambda_{T-i}}{1 - \beta_{T,T}} \quad (3)$$

with the convention $\sum_{i=0}^{-1} = 0$. Moreover, for $j = 1, \dots, T$ and $h \geq 0$, $\gamma_{j+kT}(h) = \gamma_j(h) = \beta_{j+h,h} \mu_j$ and $\gamma_{j+kT}(-h) = \gamma_j(-h) = \beta_{j+kT,h} \mu_{j+kT-h}$.

Note that the mean μ_j can be calculated recursively through the expression

$$\mu_j = \beta_{j,j}(\mu_T + \frac{1}{\beta_{j,j}} \sum_{k=0}^{j-1} \beta_{j,k} \lambda_{j-k}), \quad j = 1, \dots, T.$$

Moreover, in contrast to the autocovariance function of a stationary series, $\gamma_j(\cdot)$ is not symmetric in h ; however $\gamma_t(-h) = \gamma_{t-h}(h)$ and $\gamma_t(h) = \gamma_{t+h}(-h)$. Furthermore, in view of the fact that h can be rewritten in the form $h = i + mT$, for some $i \in \{1, \dots, T\}$ and $m \in \mathbb{N}$, the autocovariance function takes the form $\gamma_j(h) = \beta_{T,T}^m \beta_{j+i,i} \mu_j$ and $\gamma_j(-h) = \beta_{T,T}^m \beta_{j+T,i} \mu_{j+T-i}$. The main result of this section is given in Theorem 2.1 below.

Theorem 2.1. *The marginal distribution of (X_t) , with $t = j + kT$ for a fixed value of $j = 1, \dots, T$, with $T \in \mathbb{N}$, and $k \in \mathbb{N}_0$, is Poisson with mean μ_j if and only if (Z_t) forms a sequence of independent Poisson-distributed random variables with mean λ_j .*

Proof. See Appendix A. □

3 Parameters estimation

Let (X_1, \dots, X_{NT}) be a sequence of r.v.'s satisfying (1) being $\boldsymbol{\theta} = (\alpha_1, \lambda_1, \dots, \alpha_T, \lambda_T)$ the vector of unknown parameters. The notation used assumes N complete cycles of observations. Without lost of generality we assume that $X_0 = x_0$. The methods of estimation under analysis in this section are the method of moments, least squares-type and likelihood-based ones.

3.1 Moments-based estimators

In this section we discuss Yule-Walker estimators (YW) for the vector of parameters $\boldsymbol{\theta}$ which consist in the solution to the Yule-Walker type equations

$$\begin{cases} \mu_i = \alpha_i \mu_{i-1} + \lambda_i \\ \gamma_{i+kT}(1) = \beta_{i+1,1} \sigma_i^2 \end{cases}, \quad (4)$$

for $i = 1, \dots, T$ and $\mu_0 = \mu_T$, with $\gamma_{i+kT}(1)$, μ_i , and σ_i^2 replaced by their corresponding empirical counterparts. Hence, the YW-estimators, $\hat{\alpha}_{i,YW}$ and $\hat{\lambda}_{i,YW}$, of α_i and λ_i , for $i = 1, \dots, T$ are

given by

$$\hat{\alpha}_{i,YW} = \begin{cases} \frac{N \sum_{k=0}^{N-1} (X_{kT} - \bar{X}_{T,N})(X_{1+nT} - \bar{X}_{1,N})}{(N-1) \sum_{k=0}^N (X_{kT} - \bar{X}_{T,N})^2}, & i = 1 \\ \frac{\sum_{k=0}^{N-1} (X_{i-1+kT} - \bar{X}_{i-1,N})(X_{i+kT} - \bar{X}_{i,N})}{\sum_{k=0}^{N-1} (X_{i-1+kT} - \bar{X}_{i-1,N})^2}, & i = 2, \dots, T \end{cases}$$

and

$$\hat{\lambda}_{i,YW} = \begin{cases} \bar{X}_{i,N} - \hat{\alpha}_i \bar{X}_{i-1,N}, & i = 2, \dots, T \\ \bar{X}_1 - \hat{\alpha}_1 \bar{X}_{T,N}, & i = 1 \end{cases}$$

with

$$\bar{X}_{i,N} = \begin{cases} \frac{1}{N} \sum_{k=0}^{N-1} X_{i+kT} & i = 1, \dots, T-1 \\ \frac{1}{N+1} \sum_{k=0}^N X_{kT} & i = T \end{cases}.$$

3.2 Conditional Least Squares Estimators (CLS)

The CLS-estimators $\hat{\boldsymbol{\theta}}_{CLS} = (\hat{\alpha}_{1,CLS}, \hat{\lambda}_{1,CLS}, \dots, \hat{\alpha}_{T,CLS}, \hat{\lambda}_{T,CLS})$ of $\boldsymbol{\theta}$ are obtained by minimizing the expression

$$Q(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \sum_{i=1}^T (X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i)^2,$$

yielding

$$\hat{\alpha}_{i,CLS} = \frac{N \sum_{k=0}^{N-1} X_{i+kT} X_{i-1+kT} - \sum_{k=0}^{N-1} X_{i+kT} \sum_{k=0}^{N-1} X_{i-1+kT}}{N \sum_{k=0}^{N-1} X_{i-1+kT}^2 - \left(\sum_{k=0}^{N-1} X_{i-1+kT} \right)^2}$$

and

$$\hat{\lambda}_{i,CLS} = N^{-1} \left(\sum_{k=0}^{N-1} X_{i+kT} - \hat{\alpha}_{i,CLS} \sum_{k=0}^{N-1} X_{i-1+kT} \right),$$

for $i = 1, \dots, T$. It is easy to check that $\hat{\alpha}_{i,YW} = \hat{\alpha}_{i,CLS}$, for $i = 2, \dots, T$ whereas $\hat{\lambda}_{i,YW} = \hat{\lambda}_{i,CLS}$, for $i = 2, \dots, T - 1$. However, it follows from straightforward modifications of the proof of Theorem 3 in Freeland and McCabe (2005) that $\hat{\alpha}_{1,CLS} = \hat{\alpha}_{1,YW} + o_p(1)$, $\hat{\lambda}_{1,CLS} = \hat{\lambda}_{1,YW} + o_p(1)$ and $\hat{\lambda}_{T,CLS} = \hat{\lambda}_{T,YW} + o_p(1)$. In obtaining the asymptotic distribution of $\hat{\boldsymbol{\theta}}_{CLS}$, we first prove the following lemma.

Lemma 3.1. *Let $S_N(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta})$ with*

$$\xi_k(\boldsymbol{\theta}) = \sum_{i=1}^T U_{i+kT} \left(\frac{\partial U_{i+kT}}{\partial \boldsymbol{\theta}} \right),$$

being $U_{i+kT} = X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i$. As $N \rightarrow \infty$

$$N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \xrightarrow{P} A_1 \quad (5)$$

with

$$A_1 = \begin{bmatrix} \Psi_1 & \boldsymbol{o}_2 & \boldsymbol{o}_2 & \dots & \boldsymbol{o}_2 \\ \boldsymbol{o}_2 & \Psi_2 & \boldsymbol{o}_2 & \dots & \boldsymbol{o}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{o}_2 & \boldsymbol{o}_2 & \boldsymbol{o}_2 & \dots & \Psi_T \end{bmatrix},$$

being $\Psi_i = \begin{bmatrix} \mu_{i-1} + \mu_{i-1}^2 & \mu_{i-1} \\ \mu_{i-1} & 1 \end{bmatrix}$ for $i=1, \dots, T$, with $\mu_0 = \mu_T$. Moreover

$$N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' \xrightarrow{P} A_2 \quad (6)$$

with

$$A_2 = \begin{bmatrix} \Omega_1 & \boldsymbol{o}_2 & \boldsymbol{o}_2 & \dots & \boldsymbol{o}_2 \\ \boldsymbol{o}_2 & \Omega_2 & \boldsymbol{o}_2 & \dots & \boldsymbol{o}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{o}_2 & \boldsymbol{o}_2 & \boldsymbol{o}_2 & \dots & \Omega_T \end{bmatrix},$$

being

$$\Omega_i = \begin{bmatrix} \alpha_i(1 - \alpha_i)m_{i-1,3} + \lambda_i m_{i-1,2} & \alpha_i(1 - \alpha_i)m_{i-1,2} + \lambda_i \mu_{i-1} \\ \alpha_i(1 - \alpha_i)m_{i-1,2} + \lambda_i \mu_{i-1} & \alpha_i(1 - \alpha_i)\mu_{i-1} + \lambda_i \end{bmatrix},$$

and $m_{i,j} = E[X_{i+kT}^j]$, $m_{0,j} = E[X_{kT}^j]$, $\alpha_0 = \alpha_T$ and $\lambda_0 = \lambda_T$, for $i = 1, \dots, T$. Finally

$$N^{-1}S_N(\boldsymbol{\theta}) \xrightarrow{P} \vec{0} \quad (7)$$

with $\vec{0} = (0, \dots, 0)'$ of dimension $2T \times 1$.

Proof. See Appendix A. □

The following result establishes the asymptotic distribution of the CLS-estimators.

Theorem 3.1. *Under the conditions above*

$$N^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}) \xrightarrow{D} N(\vec{0}, A_1^{-1}A_2A_1^{-1}).$$

Proof. See Appendix A. □

3.3 Weighted Conditional Least Squares Estimators (WCLS)

The WCLS-estimators $\hat{\boldsymbol{\theta}}_{WCLS} = (\hat{\alpha}_{1,WCLS}, \hat{\lambda}_{1,WCLS}, \dots, \hat{\alpha}_{T,WCLS}, \hat{\lambda}_{T,WCLS})$ of $\boldsymbol{\theta}$ are obtained by minimizing the sum of the squared error between the observations and its conditional mean and weighted by the inverse of the conditional variance, given by

$$Q(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \sum_{i=1}^T \frac{(X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i)^2}{\hat{\alpha}_i(1 - \hat{\alpha}_i)X_{i+kT-1} + \hat{\lambda}_i},$$

yielding

$$\begin{aligned} \hat{\alpha}_{i,WCLS} &= \frac{\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} \sum_{k=0}^{N-1} \frac{X_{i+kT} X_{i+kT-1}}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} - \sum_{k=0}^{N-1} \frac{X_{i+kT}}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} \sum_{k=0}^{N-1} \frac{X_{i+kT-1}}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})}}{\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} \sum_{k=0}^{N-1} \frac{X_{i+kT-1}^2}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} - \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})} \right)^2}, \\ \hat{\lambda}_{i,WCLS} &= \frac{\sum_{k=0}^{N-1} X_{i+kT} - \hat{\alpha}_{i,WCLS} \sum_{k=0}^{N-1} X_{i+kT-1}}{\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1})}}, \end{aligned}$$

being $\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1}) = \hat{\alpha}_{i,CLS}(1 - \hat{\alpha}_{i,CLS})X_{i+kT-1} + \hat{\lambda}_{i,CLS}$, for $i = 1, \dots, T$.

3.4 Conditional Maximum Likelihood Estimation (CML)

The conditional likelihood function for the PINAR(1)_T model can be shown to be

$$L(\boldsymbol{\theta}|X_1, \dots, X_{NT}) = \prod_{k=0}^{N-1} \prod_{i=1}^T p_i(x_{i+kT}|x_{i-1+kT})$$

with

$$p_i(x_{i+kT}|x_{i-1+kT}) = e^{-\lambda_i} \sum_{m=0}^{\min(x_{i+kT}, x_{i-1+kT})} C_m^{x_{i-1+kT}} \alpha_i^m (1 - \alpha_i)^{x_{i-1+kT}-m} \frac{\lambda_i^{x_{i+kT}-m}}{(x_{i+kT} - m)!}.$$

The CML-estimators $\hat{\boldsymbol{\theta}}_{CML} = (\hat{\alpha}_{1,CML}, \hat{\lambda}_{1,CML}, \dots, \hat{\alpha}_{T,CML}, \hat{\lambda}_{T,CML})$ are obtained maximizing the conditional log-likelihood function

$$l(\boldsymbol{\theta}|X_1, \dots, X_{NT}) = \sum_{k=0}^{N-1} \sum_{i=1}^T \ln(p_i(x_{i+kT}|x_{i-1+kT})),$$

yielding the systems

$$\begin{cases} \sum_{k=0}^{N-1} (x_{i+kT} - \alpha_i x_{i-1+kT}) &= \lambda_i \sum_{k=0}^{N-1} \frac{p_i(x_{i+kT} - 1|x_{i-1+kT})}{p_i(x_{i+kT}|x_{i-1+kT})} \\ \sum_{k=0}^{N-1} \frac{p_i(x_{i+kT} - 1|x_{i-1+kT})}{p_i(x_{i+kT}|x_{i-1+kT})} &= N \end{cases}, \quad (8)$$

for $i = 1, \dots, T$. In order to solve this system, numerical procedures have to be employed. Note, however, that the CML estimates for the λ_i 's are readily available from that for the α_i 's through the following expression

$$\hat{\lambda}_{i,CML} = \frac{1}{N} \sum_{k=0}^{N-1} (x_{i+kT} - \hat{\alpha}_{i,CML} x_{i-1+kT}).$$

Large sample distribution of the CLM-estimators is given below.

Theorem 3.2. *The CML-estimators are asymptotically Normal, i.e.*

$$\sqrt{N} \begin{bmatrix} \hat{\alpha}_{1,CML} - \alpha_1 \\ \hat{\lambda}_{1,CML} - \lambda_1 \\ \vdots \\ \hat{\alpha}_{T,CML} - \alpha_T \\ \hat{\lambda}_{T,CML} - \lambda_T \end{bmatrix} \xrightarrow{D} N(\vec{0}, I^{-1}), \quad (9)$$

being

$$I = \begin{bmatrix} M_1 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & M_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & M_T \end{bmatrix}, \quad M_i = \begin{bmatrix} -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i^2}\right] & -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] \\ -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] & -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda_i^2}\right] \end{bmatrix}$$

with

$$\begin{aligned} -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i^2}\right] &= \\ &= -\frac{N}{\alpha_i^2(1-\alpha_i)^2} \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \\ &\times \left\{ [(2\alpha_i - 1)x_{i+T} - \alpha_i^2 x_{i-1+T}] p_i(x_{i+T} | x_{i-1+T}) + \right. \\ &+ \left. 2(1 - \alpha_i) \lambda_i p_i(x_{i+T} - 1 | x_{i-1+T}) + \lambda_i^2 p_i(x_{i+T} - 2 | x_{i-1+T}) - \lambda_i^2 \frac{p_i^2(x_{i+T} - 1 | x_{i-1+T})}{p_i(x_{i+T} | x_{i-1+T})} \right\}, \end{aligned}$$

$$\begin{aligned} -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] &= \\ &= \frac{N}{\alpha_i(1-\alpha_i)} \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \\ &\times \left\{ p_i(x_{i+T} - 1 | x_{i-1+T}) + \lambda_i p_i(x_{i+T} - 2 | x_{i-1+T}) - \lambda_i \frac{p_i^2(x_{i+T} - 1 | x_{i-1+T})}{p_i(x_{i+T} | x_{i-1+T})} \right\} \end{aligned}$$

and

$$\begin{aligned} -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda_i^2}\right] &= \\ &= -N \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \left\{ p_i(x_{i+T} - 2 | x_{i-1+T}) - \frac{p_i^2(x_{i+T} - 1 | x_{i-1+T})}{p_i(x_{i+T} | x_{i-1+T})} \right\}. \end{aligned}$$

Proof. See Appendix A. □

4 Simulation study

The aim of this section is to illustrate the theoretical findings given in Section 3 and to assess the small, moderate and large sample behavior of the YW, CLS, WCLS and CML estimators.

Throughout the analysis it shall be assumed that $T = 4$. The simulation study contemplates the following combination of λ 's and α 's: $\boldsymbol{\alpha} = (0.85, 0.50, 0.76, 0.63)$ and $\boldsymbol{\lambda} = (4, 1, 3, 2)$. We simulated time series of length $n = NT = 80, 400, 2000$ with 1000 independent replicates. The results are summarized in Table 1.

(Table 1 about here)

A closer look at the table shows the superiority of the CML method in terms of both bias and mean square error (MSE) when the sample size increases. The simulation study also reveals that there is no gain in considering the WCLS method over the CLS method, being the differences between their sample means and their mean square errors negligible. Note that the YW, CLS and WCLS estimates of $\boldsymbol{\alpha}$, componentwise, tend to be biased to the left and negatively skewed; see Figure 3. The bias and skewness, however, are reduced when the sample size increases. This is in contrast with the YW, CLS and WCLS estimates of $\boldsymbol{\lambda}$ which tend to be biased to the right and positively skewed, specially for the largest values of the λ 's. As expected, both the bias and the skewness are also reduced when the sample size increases; see Figure 4. This is in agreement with the asymptotic properties of the estimators: unbiasedness and consistency.

(Figure 3 about here)

(Figure 4 about here)

5 Conclusions

This paper has presented the periodic integer-valued autoregressive model of order one with period T , driven by a periodic sequence of independent Poisson-distributed random variables, by extending the conventional periodic autoregressive model. The stationarity and ergodicity of the process are established. Method of moments, least squares-type and likelihood-based estimators of the model parameters are derived and their asymptotic properties obtained. Moreover, an important issue in fitting conventional PARMA models to periodic time series lies with parsimony. Even very simple PARMA models can have an inordinately large number of parameters. This is also true when dealing with PINAR models. Therefore, the developments of procedures for dimensionality reduction is an impeding problem. This remains a topic of future research.

6 Appendix A

Proof of Proposition 2.1. It is easy to see that X_t is a Markov chain on \mathbb{N}_0 with the following transition probabilities

$$P_{ab}^{(1)} = P(X_t = b | X_{t-1} = a) = \sum_{m=0}^{\min(b,a)} \phi_t^m (1 - \phi_t)^{a-m} e^{-v_t} \frac{v_t^{b-m}}{(b-m)!} > 0. \quad (10)$$

From (10) it follows that (X_t) is an irreducible and aperiodic Markov chain. Thus, it is either positive recurrent or $\lim_{k \rightarrow +\infty} P_{ab}^{(j+kT)} = 0$, for any $a, b \in \mathbb{N}_0$. By iterating equation (1) and after rearranging some terms, it follows that

$$\begin{aligned} X_{j+kT} &\stackrel{d}{=} \beta_{j+kT, j+kT} \circ X_0 + \sum_{m=1}^k \sum_{a=0}^{T-1} (\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1}) \circ Z_{T(m+1)-a} + \sum_{m=0}^{j-1} \beta_{j,m} \circ Z_{j-m} \quad (11) \\ &= Y_{j+kT}, \text{ say.} \end{aligned}$$

The equality in distribution holds both unconditionally and conditionally given $X_0 = x_0$, say. Note that the first term on the right-hand side of (11) is $o_p(1)$, both unconditionally and conditionally.

Next we show that Y_{j+kT} converges almost surely. For all $\epsilon > 0$ and $k, n \in \mathbb{N}_0$

$$\begin{aligned} &P \left(\max_{1 \leq l \leq n} |Y_{j+kT} - Y_{j+T(k+l)}| > \epsilon \right) \\ &\leq \epsilon^{-1} \cdot E \left[\beta_{j,j} \beta_{T,T}^k \circ X_0 - \beta_{j,j} \beta_{T,T}^{k+n} \circ X_0 + \sum_{m=1}^n \sum_{a=0}^{T-1} (\beta_{j,a} \beta_{T,a} \beta_{T,T}^{m+k-1}) \circ Z_{T(m+k+1)-a} \right] \\ &= \epsilon^{-1} \cdot (\beta_{j,j} \beta_{T,T}^k (1 - \beta_{T,T}^n) E(X_0) + \beta_{j,j} \beta_{T,T}^{k-1} \sum_{m=1}^n \sum_{a=0}^{T-1} \beta_{T,a} \beta_{T,T}^m \lambda_{T-a}) \longrightarrow 0, \text{ as } k \rightarrow +\infty. \end{aligned}$$

Note that the variable $V_j^{(k)} = \sum_{m=1}^k \sum_{a=0}^{T-1} \beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1} \circ Z_{T(m+1)-a} + \sum_{m=0}^{j-1} \beta_{j,m} \circ Z_{j-m}$, converges in distribution (both unconditionally and conditionally) to the same limit. Hence, it follows that

$$\begin{aligned} \lim_{k \rightarrow +\infty} P(X_{j+kT} = b | X_0 = a) &= \lim_{k \rightarrow +\infty} P(Y_{j+kT} = b | X_0 = a) = \lim_{k \rightarrow +\infty} P(V_j^{(k)} = b) \\ &= \lim_{k \rightarrow +\infty} P(Y_{j+kT} = b) = \lim_{k \rightarrow +\infty} P(X_{j+kT} = b), \end{aligned}$$

since $V_j^{(k)}$ is independent of X_0 . Suppose now that $\lim_{k \rightarrow +\infty} P_{ab}^{(j+kT)} = 0 = P(Y_j^* = b)$, for any $a, b \in \mathbb{N}_0$, where Y_j^* represents the almost sure limit of Y_{j+kT} . However, this contradicts the fact

that $P(Y_j^* < \infty) = 1$. This proves the first part of the proposition. Almost sure convergence of $V_j^{(k)}$ follows by the same argument as above. We skip the details. Finally, in order to prove the convergence in L_2 , we proceed as follows:

$$\begin{aligned} E \left[\left(V_j^{(k)} - V_j^{(k+n)} \right)^2 \right] &= \sum_{m=1}^n \sum_{\substack{l=1 \\ l \neq m}}^n \sum_{a=0}^{T-1} \sum_{b=0}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{m+l+2k-2} \lambda_{T-a} \lambda_{T-b} \\ &+ \sum_{m=1}^n \sum_{a=0}^{T-1} \left(\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m+k-1} \lambda_{T-a} + \beta_{j,j}^2 \beta_{T,a}^2 \beta_{T,T}^{2m+2k-2} \lambda_{T-a}^2 \right) \\ &+ \sum_{m=1}^n \sum_{a=0}^{T-1} \sum_{\substack{b=0 \\ b \neq a}}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{2m+2k-2} \lambda_{T-a} \lambda_{T-b}. \end{aligned}$$

After some tedious calculations, we obtain

$$\begin{aligned} E \left[\left(V_{j+kT} - V_{j+T(k+n)} \right)^2 \right] &= \beta_{T,T}^{2k+1} \beta_{j,j}^2 \times (\text{constant}) + \beta_{T,T}^k \beta_{j,j} \times (\text{constant}) \\ &+ \beta_{T,T}^{2k} \beta_{j,j}^2 \times (\text{constant}) \longrightarrow 0, \quad k \rightarrow +\infty. \end{aligned}$$

Therefore, for all $n > 0$, $E \left[\left| V_{j+kT} - V_{j+T(k+n)} \right|^2 \right] \longrightarrow 0$, as $k \rightarrow +\infty$. This concludes the proof.

Proof of Theorem 2.1. We first prove the *if* part: assume that $Z_{j+kT} \sim P(\lambda_j)$. Note that X_{j+kT} can be rewritten as

$$X_{j+kT} = \beta_{T,T} \circ X_{j+T(k-1)} + \underbrace{\sum_{i=0}^{j-1} \beta_{j,i} \circ Z_{j-i} + \sum_{i=0}^{T-j-1} \beta_{j,j} \beta_{T,i} \circ Z_{T-i}}_{Y_j^*}.$$

Since $\sum_{m=0}^{kT-1} \beta_{j+kT,m} \circ Z_{j+kT-m} \sim P((1 - \beta_{T,T}^k) \mu_j)$, it follows that $Y_j^* \sim P((1 - \beta_{T,T}) \mu_j)$. Moreover, the probability generating function of X_{j+kT} takes the form $P_{X_{j+kT}}(r) = P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T} r) \times P_Y(r)$ which implies that

$$\frac{P_{X_{j+kT}}(r)}{P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T} r)} = e^{-\mu_j(1 - \beta_{T,T})(1-r)} = \frac{e^{-\mu_j(1-r)}}{e^{-\mu_j(1 - (1 - \beta_{T,T} + \beta_{T,T} r))}},$$

which lead us to conclude that $P_{X_{j+kT}}(r) = e^{-\mu_j(1-r)}$ is the probability generating function of a Poisson-distributed random variable of mean μ_j . Finally, we prove the *only if* part. If X_{j+kT} is a Poisson random variable, then

$$e^{-\mu_j(1-r)} = P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T} r) \times P_Y(r) = e^{-\mu_j \beta_{T,T}(1-r)} \times P_Y(r),$$

providing that

$$P_Y(r) = \prod_{i=0}^{j-1} e^{-\beta_{j,i}\lambda_{j-i}(1-r)} \prod_{i=0}^{T-j-1} e^{-\beta_{j,j}\beta_{T,i}\lambda_{T-i}(1-r)} \quad (12)$$

$$= \prod_{i=0}^{j-1} P_{Z_{j-i}}(1 - \beta_{j,i} + \beta_{j,i}r) \prod_{i=0}^{T-j-1} P_{Z_{T-i}}(1 - \beta_{j,j}\beta_{T,i} + \beta_{j,j}\beta_{T,i}r), \quad (13)$$

which is the probability generating function of a Poisson r.v. of mean $\mu_j^* = \sum_{m=0}^{j-1} \beta_{j,m}\lambda_{j-m} + \beta_{j,j} \sum_{m=0}^{T-j-1} \beta_{T,m}\lambda_{T-m}$. By (12) and (13) it follows that

$$P_{Z_{j-i}}(1 - \beta_{j,i} + \beta_{j,i}r) = e^{-\lambda_{j-i}(1-\beta_{j,i}+\beta_{j,i}r)}$$

and

$$P_{Z_{T-i}}(1 - \beta_{j,j}\beta_{T,i} + \beta_{j,j}\beta_{T,i}r) = e^{-\lambda_{T-i}(1-\beta_{j,j}\beta_{T,i}+\beta_{j,j}\beta_{T,i}r)}.$$

Thus, $P_{Z_{j-i}}(s) = e^{-\lambda_{j-i}(1-s)}$ and $P_{Z_{T-i}}(r) = e^{-\lambda_{T-i}(1-r)}$ which implies $Z_i \sim P(\lambda_i)$, for $i = 1, \dots, T$. This completes the proof.

Proof of Lemma 3.1. First, note that $E[N^{-1}\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}] = A_1$. Thus the proof of (5) is concluded if we show that

$$V[N^{-1}\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}] = E\left[N^{-2}\left(\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^2\right] - A_1^2 \rightarrow \mathbf{0}_{2T}. \quad (14)$$

In proving (14) observe that

$$N^{-2}\left(\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^2 = \begin{bmatrix} M_{1,N} & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & M_{2,N} & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & M_{s,N} \end{bmatrix}$$

with

$$M_{i,N} = N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} \begin{bmatrix} X_{i-1+nT}^2 X_{i-1+kT}^2 + X_{i-1+nT} X_{i-1+sk} & X_{i-1+sn}^2 X_{i-1+sk} + X_{i-1+sn} \\ X_{i-1+nT} X_{i-1+kT}^2 + X_{i-1+kT} & N^{-2} X_{i-1+nT} X_{i-1+kT} + 1 \end{bmatrix}.$$

After some tedious but straightforward calculations it is possible to prove that for $r, w = 0, 1, 2$

$$\begin{aligned} \lim_{N \rightarrow +\infty} N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E[X_{i-1+nT}^r X_{i-1+kT}^w] &= \frac{1}{2} (I_{\{r=0\}} + I_{\{r=1\}} m_{i-1,1} + I_{\{r=2\}} m_{i-1,2}) m_{i-1,w} \\ &+ \frac{1}{2} (I_{\{w=0\}} + I_{\{w=1\}} m_{i-1,1} + I_{\{w=2\}} m_{i-1,2}) m_{i-1,r}. \end{aligned}$$

Hence,

$$\lim_{N \rightarrow +\infty} E[M_{i,N}] = \Psi_i^2 = \begin{bmatrix} m_{i-1,2}^2 + m_{i-1,1}^2 & m_{i-1,2}m_{i-1,1} + m_{i-1,1} \\ m_{i-1,2}m_{i-1,1} + m_{i-1,1} & m_{i-1,1}^2 + 1 \end{bmatrix},$$

for $i = 1, \dots, T$, providing

$$\lim_{N \rightarrow +\infty} E \left[\left(N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - A_1 \right) \left(N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - A_1 \right)^T \right] = \mathbf{0}_{2T}.$$

This proves the result in (14). To prove (6) we proceed as follows: by symmetry of the matrix in (6)

$$N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' = N^{-1} \sum_{k=0}^{N-1} \begin{bmatrix} C_{k11} & C_{k12} & C_{k13} & \dots & C_{k1T} \\ C_{k12} & C_{k22} & C_{k23} & \dots & C_{k2T} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{k1T} & C_{k2T} & C_{k3T} & \dots & C_{kTT} \end{bmatrix} \quad (15)$$

with

$$C_{kij} = \begin{bmatrix} U_{i+kT} X_{i-1+kT} U_{j+kT} X_{j-1+kT} & U_{i+kT} U_{j+kT} X_{j-1+kT} \\ U_{i+kT} X_{i-1+kT} U_{j+kT} & U_{i+kT} U_{j+kT} \end{bmatrix}, j \geq i.$$

Note that the (m, l) th entry of the matrix C_{kij} , denoted by $(C_{kij})_{ml}$, for $j, i = 1, \dots, T, j \neq i$ and $l, m = 1, 2$, can be expressed in the form $(C_{kij})_{ml} = U_{i+kT} X_{i-1+kT}^{2-l} U_{j+kT} X_{j-1+kT}^{2-m}$ whereas $(C_{kii})_{ml} = U_{j+kT}^2 X_{i-1+kT}^{4-m-l}$. Consequently,

$$E[(C_{Nij})_{ml}] = E[U_{i+kT} X_{i-1+kT}^{2-l} X_{j-1+kT}^{2-m} \underbrace{E[U_{j+kT} | X_{i-1+kT}, U_{i+kT}, X_{j-1+kT}]}_0] = 0$$

and

$$\begin{aligned} E[(C_{Nii})_{ml}] &= E[X_{i-1+kT}^{4-m-l} \underbrace{E[(X_{i+kT} - \alpha_i X_{i-1+kT} - \lambda_i)^2 | X_{i-1+kT}]}_{V[X_{i+kT} | X_{i-1+kT}]}] \\ &= \alpha_i(1 - \alpha_i)m_{i-1, (5-m-l)} + \lambda_i m_{i-1, (4-m-l)}. \end{aligned}$$

Thus, it follows that $E[N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})'] = A_2$. Again, after some tedious but straightforward calculations it can be prove that

$$E \left[\left(N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' - A_2 \right) \left(N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' - A_2 \right)' \right] \rightarrow \mathbf{0}_{2T},$$

as $N \rightarrow \infty$. This proves the result in (6). The proof of Lemma 3.1 is complete upon showing the result in (7). In doing so, first note that $E[U_{i+kT}] = 0$ and $E[U_{i+kT}X_{i-1+kT}] = E[X_{i-1+kT}E[U_{i+kT}|X_{i-1+kT}]] = 0$, for $i = 1, \dots, T$, leading to

$$E[N^{-1}S_N(\boldsymbol{\theta})] = \vec{0}.$$

In order to calculate the variance-covariance matrix of $N^{-1}S_N(\boldsymbol{\theta})$, the following preliminary results are needed

$$N^{-1}S_N(\boldsymbol{\theta}) (N^{-1}S_N(\boldsymbol{\theta}))' = N^{-2} \sum_{n=0}^{N-1} \sum_{\substack{k=0 \\ k \neq n}}^{N-1} E[\xi_n(\boldsymbol{\theta})\xi_k(\boldsymbol{\theta})'] + N^{-2} \sum_{n=0}^{N-1} \xi_n(\boldsymbol{\theta})\xi_n(\boldsymbol{\theta})'$$

with

$$\sum_{n=0}^{N-1} \sum_{\substack{k=0 \\ k \neq n}}^{N-1} \xi_n(\boldsymbol{\theta})\xi_k(\boldsymbol{\theta})' = \sum_{n=0}^{N-1} \sum_{\substack{k=0 \\ k \neq n}}^{N-1} \begin{bmatrix} F_{1,1} & F_{1,2} & F_{1,3} & \dots & F_{1,T} \\ F_{2,1} & F_{2,2} & F_{2,3} & \dots & F_{2,T} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ F_{T,1} & F_{T,2} & F_{T,3} & \dots & F_{T,T} \end{bmatrix},$$

being $F_{i,j}$, for $i, j = 1, \dots, T$, zero-mean (2×2) matrices with (m, l) th entry given by

$$(F_{i,j})_{ml} = U_{i+nT}X_{i-1+nT}^{2-l}U_{j+kT}X_{j-1+kT}^{2-m}, \quad m, l = 1, 2.$$

Thus, the variance-covariance matrix of $N^{-1}S_N(\boldsymbol{\theta})$ takes the form

$$E[N^{-1}S_N(\boldsymbol{\theta}) (N^{-1}S_N(\boldsymbol{\theta}))'] = N^{-2} \sum_{n=0}^{N-1} E[\xi_n(\boldsymbol{\theta})\xi_n(\boldsymbol{\theta})'] = N^{-1}A_2 \rightarrow \mathbf{0}_{2T},$$

as $N \rightarrow \infty$. The proof is complete.

Proof of Theorem 3.1. Let $\mathcal{F}_{k-1} = \sigma(X_{kT}, X_{(k-1)T}, \dots)$. Since $E[\xi_k(\boldsymbol{\theta})|\mathcal{F}_{k-1}] = \vec{0}$ it turns out that $\xi_k(\boldsymbol{\theta})$ is a zero-mean martingale with respect to \mathcal{F}_k . By (6) it follows that the central limit theorem for martingales (Hall and Heyde, 1980) applies to $S_k(\boldsymbol{\theta})$, providing

$$N^{-\frac{1}{2}}S_N(\boldsymbol{\theta}) \xrightarrow{D} N(\vec{0}, A_2), \quad N \rightarrow \infty. \quad (16)$$

Now consider the first-order Taylor expansion

$$S_N(\hat{\boldsymbol{\theta}}_{CLS}) = S_N(\boldsymbol{\theta}) + \left(\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) (\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}) + R_N, \quad (17)$$

being the remainder term $R_N = O_p(1)$, as $N \rightarrow \infty$. Now use $S_N(\hat{\boldsymbol{\theta}}_{CLS}) = 0$ in (17) to get

$$N^{1/2}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}) = - \left(N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^{-1} N^{-1/2} S_N(\boldsymbol{\theta}) + o_p(1). \quad (18)$$

By (5) in Lemma 3.1, equation (16) and Slutsky's Theorem in equation (18) the result follows.

Proof of Theorem 3.2. The proof follows easily as a generalization, for the periodic case, of the arguments given by Franke and Seligmann (1993, pp. 325-6). We omit the details.

Acknowledgments

The authors would like to express their gratitude to an anonymous referee for all helpful comments and constructive criticism.

References

- [1] Ahn, S., Gyemin, L., Jongwoo, J., 2000. Analysis of the M/D/1-type queue based on an integer-valued autoregressive process. *Oper. Res. Lett.* 27, 235–241.
- [2] Basawa, I.V., Lund, R.B., 2001. Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.* 22, 651–666.
- [3] Basawa, I.V., Lund, R.B., Shao, Q., 2004. First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett.* 67, 299–306.
- [4] Bennett, W.R., 1958. Statistics of regenerative digital transmission. *Bell Sys Tech. J.* 37, 1501-1542.
- [5] Bentarzi, M., Hallin, M., 1994. On the invertibility of periodic moving-average models. *J. Time Ser. Anal.* 15, 263–268.
- [6] Bloomfield, P., Hurd, H.L., Lund, R.B., 1994. Periodic correlation in stratospheric ozone data. *J. Time Ser. Anal.* 15, 127–150.
- [7] Brännäs, K., Nordström, J., 2006. Tourist accomodation effects of festivals. *Tourism Economics* 12, 291–302.

- [8] Cui, Y., Lund, R., 2009. A new look at time series of counts. *Biometrika* 96, 781–792.
- [9] Franke, J., Seligmann, T., 1993. Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In: Subba Rao, T. (Ed.), *Developments in Time Series Analysis*. Chapman & Hall, London, 310–330.
- [10] Franses, P.H., 1994. A multivariate approach to modeling univariate seasonal time series. *J. Econometrics* 63, 133–151.
- [11] Franses, P.H., Paap, R., 2004. *Periodic Time Series*. Oxford University Press. Oxford.
- [12] Freeland, R.K., McCabe, B., 2005. Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statist. Probab. Lett.* 73, 147–153.
- [13] Gardner, W.A., Napolitano, A., Paura, L., 2006. Cyclostationary: half a century of research. *Signal Process.* 86, 639–697.
- [14] Gladyshev, E.G., 1961. Periodically correlated random sequences. *Soviet Math.* 2, 385–388.
- [15] Gladyshev, E.G., 1963. Periodically and almost periodically correlated random processes with continuous time parameter. *Theory Probab. Appl.* 8, 173–177.
- [16] Haldrup, N., Hyllerberg, S., Pons, G., Sansó, A., 2007. Common periodic correlation features and the interaction of stocks and flows in daily airport data. *J. Bus. Econom. Statist.* 25, 21–32.
- [17] Hall, P., Heyde, C.C., 1980. *Martingale Limit Theory and its Applications*. Academic Press, New York.
- [18] Hipel, K.W., McLeod, A.I., 1994. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam.
- [19] Hurd, H.L., Miamee, A., 2007. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. John Wiley & Sons, New Jersey.
- [20] Latour, A., 1997. The multivariate GINAR(p) process. *Adv. Appl. Probab.* 29, 228–248.
- [21] Lund, R.B., Basawa, I.V., 2000. Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.* 21, 75–93.

- [22] Lund, R.B., Hurd, H.L., Bloomfield, P., Smith, R.L., 1995. Climatological time series with periodic correlation. *J. Climate* 11, 2787–2809.
- [23] Lund, R.B., Shao, Q., Basawa, I.V., 2006. Parsimonious periodic time series modeling. *Aust. N. Z. J. Stat.* 48, 33–47.
- [24] Lund, R.B., Wang, X.L., Lu, Q.Q., Reeves, J., Gallagher, C., Feng, Y., 2007. Changepoint detection in periodic and autocorrelated time series. *J. Climate* 11, 5178–5190.
- [25] McCabe, B.P.M., Martin, G.M., 2005. Bayesian prediction of low count time series. *Int. J. Forecast.* 21, 315–330.
- [26] Monteiro, M., Pereira, I., Scotto, M.G., 2008. Optimal alarm systems for count processes. *Comm. Statist. Theory Methods* 37, 3054–3076.
- [27] Quoreshi, A.M.M.S., 2006. Bivariate time series modelling of financial count data. *Comm. Statist. Theory Methods* 35, 1343–1358.
- [28] Salas, J.D., 1993. Analysis and modeling of hydrologic time series. In: Maidment, D.R. (Ed.), *Handbook of Hydrology*, Chapter 19. McGraw-Hill, New York.
- [29] Shao, Q., Ni, P.P., 2004. Least-squares estimation and ANOVA for periodic autoregressive time series. *Statist. Probab. Lett.* 69, 287–297.
- [30] Shao, Q., 2006. Mixture periodic autoregressive time series models. *Statist. Probab. Lett.* 76, 609–618.
- [31] Tesfaye, Y.G., Meerschaert, M.M., Anderson, P.L., 2006. Identification of PARMA models and their application to the modeling of riverflows. *Water Resour. Res.* 42, W01419 (11pp.).
- [32] Vecchia, A.V., 1985. Periodic autoregressive-moving average (PARMA) modelling with applications to water resources. *Water Res. Bull.* 21, 721–730.
- [33] Zhou, J., Basawa, I.V., 2005. Least-squared estimation for bifurcation autoregressive processes. *Statist. Probab. Lett.* 74, 77–88.
- [34] Weiß, C.H., 2008. The combined INAR(p) models for time series of counts. *Statist. Probab. Lett.* 78, 1817–1822.

- [35] Weiß, C.H., 2009. Controlling jumps in correlated processes of Poisson counts. Appl. Stochastic Models Bus. Ind. 25, 551–564.

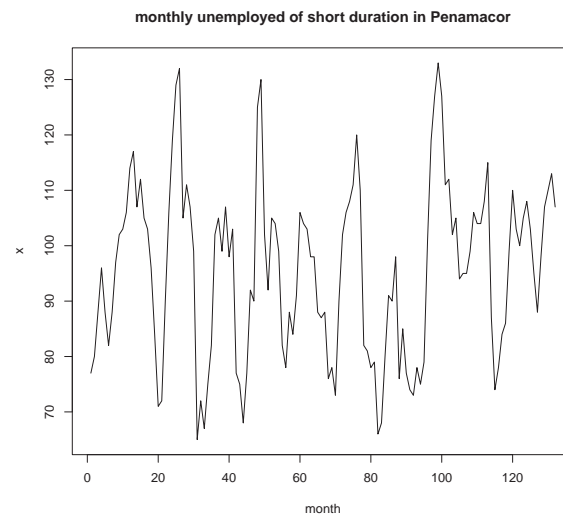


Figure 1: Monthly number of short-term unemployed people in Penamacor County (Portugal), from January 1997 to December of 2007.

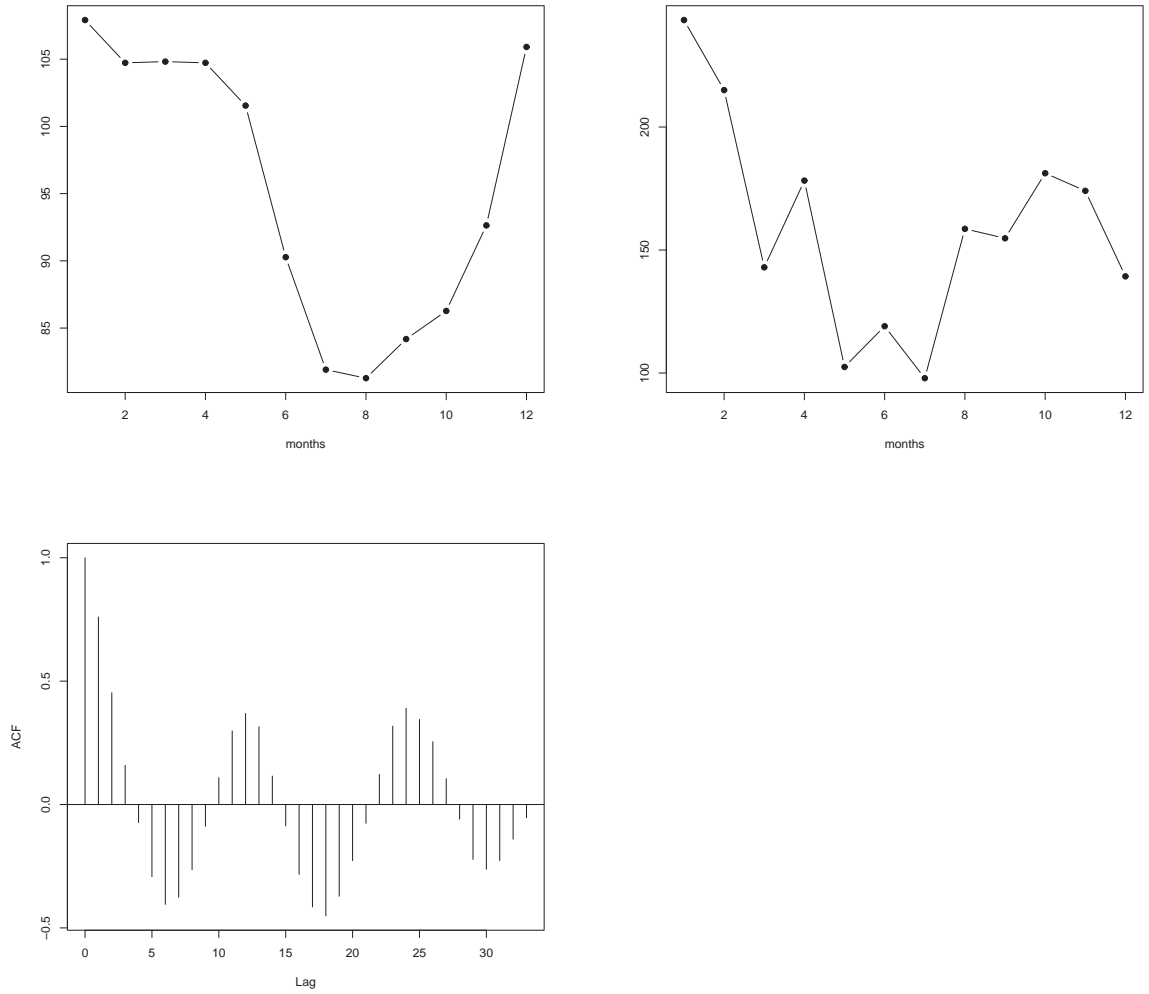


Figure 2: Sample means (top row, left) and variances (top row, right) by months of the year and the autocorrelation function (bottom row, left) of the monthly number of unemployed people in Penamacor County (Portugal), from January 1997 to December of 2007.

		α				λ			
$n = 80$	YW	0.737	0.423	0.706	0.597	4.717	1.771	3.291	2.264
		(0.039)	(0.018)	(0.036)	(0.028)	(1.965)	(1.728)	(1.356)	(1.488)
	CLS	0.740	0.423	0.706	0.597	4.697	1.771	3.291	2.255
		(0.037)	(0.018)	(0.036)	(0.028)	(1.882)	(1.728)	(1.356)	(1.488)
	WCLS	0.741	0.422	0.706	0.597	4.692	1.786	3.289	2.259
$n = 400$	YW	0.837	0.492	0.754	0.628	4.080	1.079	3.040	2.018
		(0.007)	(0.003)	(0.008)	(0.006)	(0.326)	(0.297)	(0.277)	(0.311)
	CLS	0.837	0.492	0.754	0.628	4.078	1.079	3.040	2.020
		(0.007)	(0.003)	(0.008)	(0.006)	(0.323)	(0.297)	(0.277)	(0.310)
	WCLS	0.838	0.492	0.754	0.627	4.077	1.083	3.039	2.023
$n = 2000$	YW	0.847	0.499	0.759	0.629	4.017	1.014	3.005	2.004
		(0.002)	(0.001)	(0.002)	(0.001)	(0.074)	(0.071)	(0.056)	(0.065)
	CLS	0.847	0.499	0.759	0.629	4.016	1.014	3.005	2.005
		(0.002)	(0.001)	(0.002)	(0.001)	(0.074)	(0.071)	(0.057)	(0.065)
	WCLS	0.847	0.499	0.760	0.629	4.015	1.017	3.003	2.004
$n = 2000$	CML	0.849	0.500	0.761	0.631	4.003	0.998	2.993	1.991
		(0.001)	(0.0003)	(0.001)	(0.0004)	(0.028)	(0.026)	(0.024)	(0.025)

Table 1: Sample mean and mean square error (in brackets) for the model $\alpha = (0.85, 0.50, 0.76, 0.63)$ and $\lambda = (4, 1, 3, 2)$.

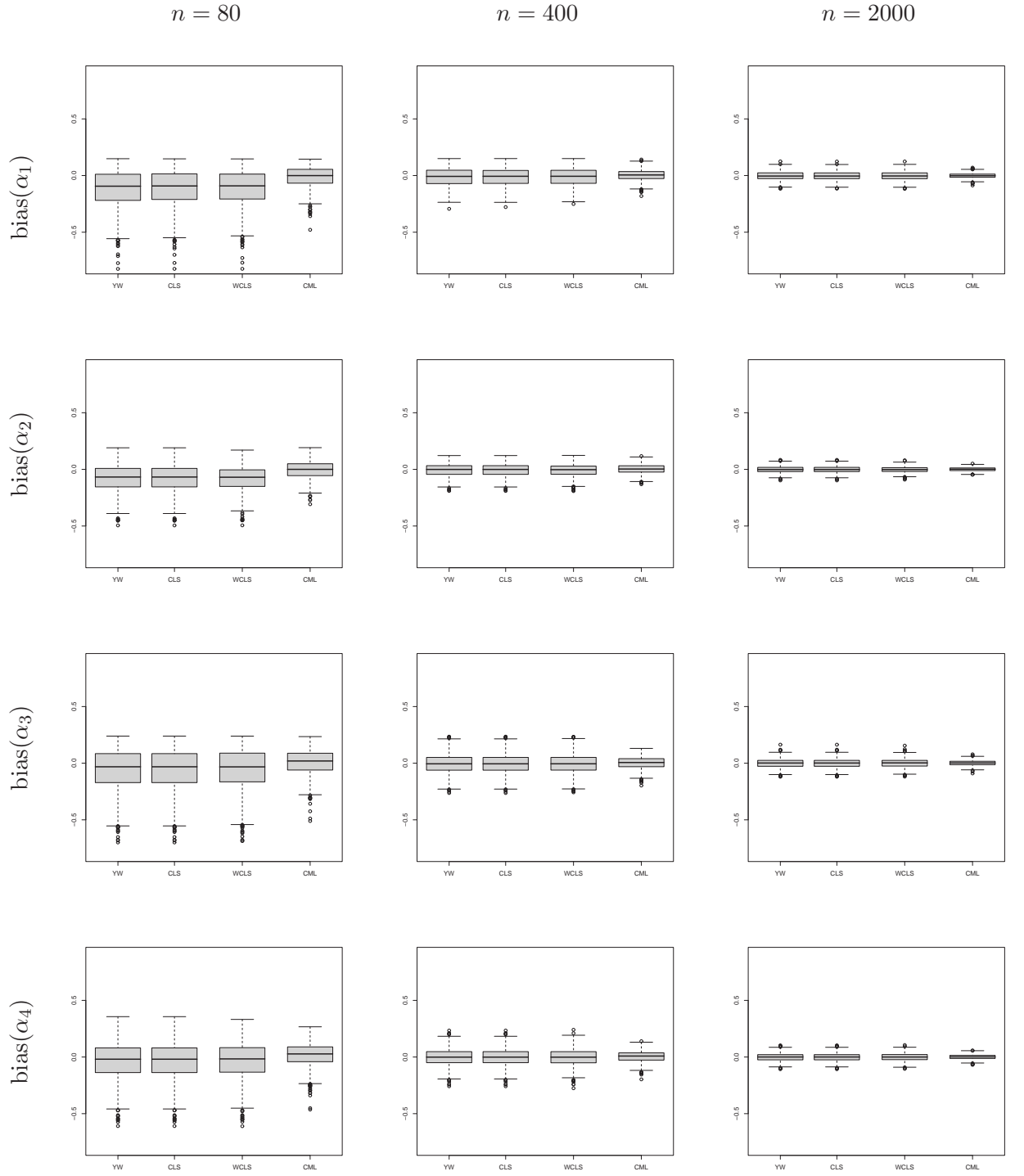


Figure 3: Boxplots of the biases for α considering the set of parameters $\alpha = (0.85, 0.50, 0.76, 0.63)$ and $\lambda = (4, 1, 3, 2)$.

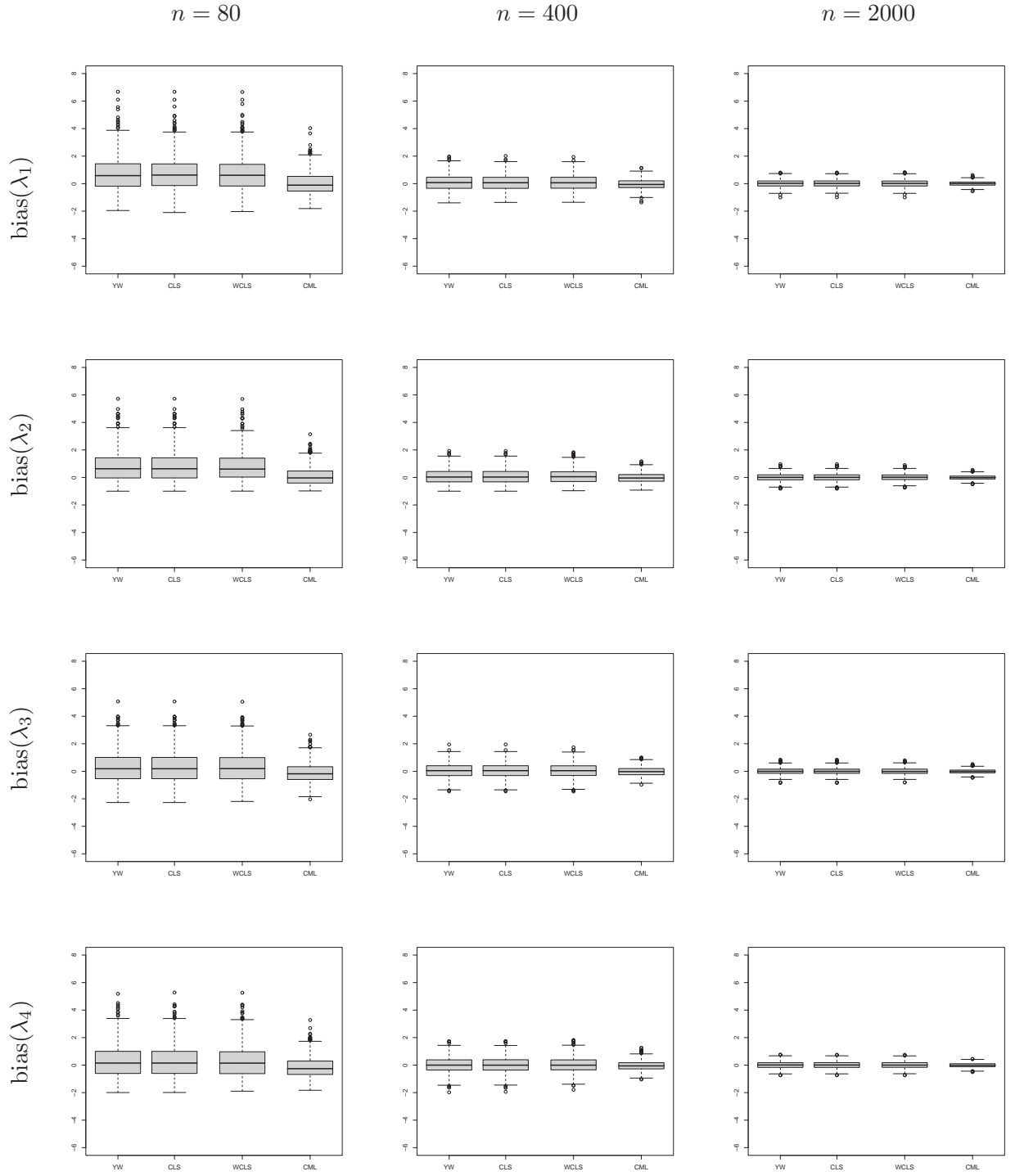


Figure 4: Boxplots of the biases for λ considering the set of parameters $\alpha = (0.85, 0.50, 0.76, 0.63)$ and $\lambda = (4, 1, 3, 2)$.